

# ON DISTRIBUTION OF THREE-TERM ARITHMETIC PROGRESSIONS IN SPARSE SUBSETS OF $\mathbf{F}_p^n$

HOI H. NGUYEN

**ABSTRACT.** We prove a structural version of Szemerédi's regularity lemma for subsets of a typical random set in  $\mathbf{F}_p^n$ . As an application, we give a short proof for an analog of a hard theorem by Kohayakawa, Łuczak, and Rödl on the distribution of three-term arithmetic progressions in sparse sets.

## 1. INTRODUCTION

Let  $G$  be a graph and let  $A, B$  be two subsets of  $V_G$ . We define the density  $d(A, B)$  of  $G(A, B)$  to be

$$d(A, B) := e(A, B)/|A||B|.$$

Let  $\varepsilon$  be a positive constant. We say that the pair  $(A, B)$  is  $\varepsilon$ -regular if

$$|d(A', B') - d(A, B)| \leq \varepsilon$$

for any  $A' \subset A$  and  $B' \subset B$  satisfying  $|A'| \geq \varepsilon|A|$  and  $|B'| \geq \varepsilon|B|$ .

Szemerédi's regularity lemma, a fundamental result in combinatorics, states that the vertex set of any dense graph can be partitioned into not-too-small pieces so that almost all pairs of pieces are regular.

**Theorem 1.1** (Szemerédi's regularity lemma). *Let  $\varepsilon > 0$ . There exists  $M = M(\varepsilon)$  such that the vertex set can be partitioned into  $1/\varepsilon \leq m \leq M$  sets  $V_i$  with sizes differing by at most 1, such that at least  $(1-\varepsilon)m^2$  of the pairs  $(V_i, V_j)$  are  $\varepsilon$ -regular.*

Consider a vector space  $V = \mathbf{F}_p^n$ , where  $p$  is a fixed odd prime and  $n$  is a large integer. Let  $A$  be a subset of  $V$ , we define the (bipartite, directed) Cayley graph generated by  $A$  to be  $G_A = G(V_1, V_2)$ , where  $V_1, V_2$  are two copies of  $V$ , and  $(v_1, v_2) \in E(G_A)$  if  $v_2 - v_1 \in A$ .

It is clear that  $G_A$  is a regular graph of degree  $|A|$ . Hence if  $A$  is dense enough, then Szemerédi's regularity lemma is applicable to  $G_A$ . Furthermore, since  $G_A$

has additional algebraic structure, it is natural to expect a stronger result than Theorem 1.1. Indeed, a result of Green [1, Section 9] confirms this intuition:

*Assume that  $|A| = \Omega(|V|)$ . Then one can partition  $V(G_A)$  into affine subspaces of large dimension and so that almost all pairs of subspaces are  $\varepsilon$ -regular.*

Szemerédi's regularity lemma is not meaningful for sparse graphs in general. However, it can be extended to certain graph families. Let  $\varepsilon$  be a positive constant. We say that the pair  $(A, B)$  is *relatively  $\varepsilon$ -regular* if

$$|d(A', B') - d(A, B)| \leq \varepsilon d(G)$$

for any  $A' \subset A$  and  $B' \subset B$  satisfying  $|A'| \geq \varepsilon|A|$  and  $|B'| \geq \varepsilon|B|$ .

Let be given  $b > 2$  and  $\sigma > 0$ . We say that a graph  $G$  is  $(b, \sigma)$ -sparse if

$$d(X, Y) \leq bd(G)$$

for any  $|X| \geq \sigma|V_G|$  and  $|Y| \geq \sigma|V_G|$ . The following result extends Szemerédi's regularity lemma for  $(b, \sigma)$ -sparse graphs.

**Theorem 1.2** (Szemerédi's regularity lemma for sparse graphs, [2, Lemma 4]). *Let  $b > 0$ . For  $\varepsilon > 0$  there exists  $\sigma = \sigma(b, \varepsilon)$  such that the following holds for all  $(b, \sigma)$ -sparse graphs. There exists  $M = M(\varepsilon, b)$  such that the vertex set can be partitioned into  $1/\varepsilon \leq m \leq M$  sets  $V_i$  with sizes differing by at most 1, such that at least  $(1 - \varepsilon)m^2$  of the pairs  $(V_i, V_j)$  are relatively  $\varepsilon$ -regular.*

As to how Theorem 1.2 extends Theorem 1.1, our first goal is to point out that the result of Green can be extended easily to “ $(b, \sigma)$ -sparse” Cayley graphs in  $\mathbf{F}_p^n$ :

*Assume that  $A$  is not too sparse, and  $G_A$  is  $(b, \sigma)$ -sparse with some reasonable constants  $b, \sigma$ . Then Theorem 1.2 is applicable to  $G_A$  in such a way that the vertex partitions can be taken to be affine subspaces of high dimension.*

We shall give a precise statement in Section 3. Next, let  $Z$  be an additive group and let  $\alpha$  be a positive constant. We say that a subset  $R$  of  $Z$  is  $(\alpha, 3AP)$ -dense if any subset of  $A$  of cardinality at least  $\alpha|A|$  must contain a nontrivial three-term arithmetic progression in  $Z$ . It has been shown in [2] that almost every subset of cardinality  $\gg_\alpha |Z|^{1/2}$  of the cyclic group  $Z = \mathbf{Z}_n$ , where  $n$  is odd, is  $(\alpha, 3AP)$ -dense. Our next goal is to prove a similar result.

**Theorem 1.3** (Kohayakawa-Łuczak-Rödl theorem for  $\mathbf{F}_p^n$ ). *There exists a constant  $C = C(\alpha)$  such that the following holds for all  $r \geq C(\alpha)|V|^{1/2}$ . Let  $R$  be a random subset of size  $r$  of  $\mathbf{F}_p^n$ , then the probability that  $R$  fails to be  $(\alpha, 3AP)$ -dense is  $o(1)$ .*

To prove Theorem 1.3 we follow the approach of [2]. However, with our structure result in hand (Theorem 3.1), we are able to get around many technical difficulties to provide a much simpler proof.

## 2. NOTATION

**Fourier transform.**(cf. [4, Chapter 4.]) Let  $H$  be a subspace of  $V$ , let  $f$  be a real-valued function defined on  $V$ . Then the Fourier transform of  $f$  with respect to  $H$  is

$$\widehat{f}(\xi) := \mathbf{E}_{x \in H} f(x) e(-\langle x, \xi \rangle)$$

Where  $\langle x, \xi \rangle = \sum_{i=1}^n x_i \xi_i / p$ , and  $e(z) = e^{2\pi i z}$ .

**Convolution.** Let  $f$  and  $g$  be two real-valued functions defined on  $V$ . The convolution of  $f$  and  $g$  with respect to  $H$  is

$$f * g(h) := \mathbf{E}_{x \in H} f(x) g(h - x).$$

The following basic properties for real-valued functions will be used several times.

- (Parseval's identity)  $\mathbf{E}_{x \in H} f^2(x) = \sum_{\xi \in H} |\widehat{f}(\xi)|^2$ .
- (Plancherel's formula)  $\mathbf{E}_{x \in H} f(x) g(x) = \sum_{\xi \in H} \widehat{f}(\xi) \overline{\widehat{g}(\xi)}$ .
- (Fourier inversion formula)  $f(x) = \sum_{\xi \in H} \widehat{f}(\xi) e(\langle x, \xi \rangle)$ .
- $\widehat{f * g}(\xi) = \widehat{f}(\xi) \widehat{g}(\xi)$ .

Let  $A$  be a subset of  $V$ , and let  $v$  be an element of  $V$ . We define  $A_H^v$  to be the set  $A + v \cap H$ . Sometimes we also write  $A_H^v$  as its characteristic function. Following are some simple properties:

- $\widehat{A_H^v}(\xi) = |A_H^v|/|H|$  if  $\xi \in H^\perp$ ;
- $\widehat{A_H^{v'}}(\xi) = e(\langle v - v', \xi \rangle) \widehat{A_H^v}(\xi)$  if  $v - v' \in H$ ; in particular,  $|\widehat{A_H^{v'}}(\xi)| = |\widehat{A_H^v}(\xi)|$ .

**$\varepsilon$ -regular vector.** Let  $\varepsilon$  be a positive constant. Let  $A$  be a given set. We say that a vector  $v$  is an  $\varepsilon$ -regular vector with respect to  $H$  if

$$\sup_{\xi \notin H^\perp} |\widehat{A_H^v}(\xi)| \leq \varepsilon |A|/|V|.$$

(It is more natural to use the upper bound  $\varepsilon|A_H^v|/|H|$  in the definition above, but we find our definition more convenient to use, and  $\varepsilon|A|/|V|$  is the typical value for  $\varepsilon|A_H^v|/|H|$ .)

Notice that if  $v$  is an  $\varepsilon$ -regular vector, then so is any element of  $v + H$ .

We say that a subspace  $H$  is  **$\varepsilon$ -regular for  $A$**  if the number of  $v$ 's which fail to be  $\varepsilon$ -regular is at most  $\varepsilon|V|$ .

**$\sigma$ -regular set.** Let  $\sigma$  be a positive constant. We say that a subset  $R$  of  $V$  is  $\sigma$ -regular if the number of edges between  $X$  and  $Y$  in the Cayley graph  $G_R$  is as many as expected,

$$e_{G_R}(X, Y) = (1 + o_\sigma(1))|R||X||Y|/|V|,$$

provided that  $|X|, |Y| \geq \sigma N$ .

Roughly speaking, a typical large random set is  $\sigma$ -regular for quite small  $\sigma$  (see Subsection 4.1). In particular, the set  $V$  itself is  $\sigma$ -regular for all  $\sigma$ .

**Dependency of constants.** We shall work with several constants throughout this note, so let us mention briefly here to avoid confusion.

$$\alpha, c(\alpha) \rightarrow \eta \rightarrow \varepsilon \rightarrow \sigma \rightarrow C$$

First,  $\alpha$  is the constant that we fix all the time. The constants  $c(\alpha)$ 's depend only on  $\alpha$ ; these constants will appear as exponents in Section 5. Next,  $\eta$  will be chosen to be small enough depending on  $\alpha$  and the  $c(\alpha)$ 's. The constant  $\varepsilon$  will be considered as an arbitrary constant in Section 3 and Subsection 4.1, but it will depend on  $\alpha$  and  $\eta$  in later sections. Last but not least,  $\sigma$  depends on  $\alpha$  and  $\varepsilon$ . We shall choose  $\eta, \varepsilon, \sigma$  to be small enough, while constants  $C = C(\alpha, \eta, \varepsilon, \sigma)$  are often very large.

**Tower-type function.** We shall use a tower-type function  $W(t)$  defined recursively by  $W(1) = 2p$  and  $W(t) = (2p)^{W(t-1)}$  for  $t \geq 2$ .

The note is organized as follows. In Section 3 we discuss about Green's result for sparse Cayley graphs. Next in Section 4 we shall provide some ingredients for applications. The proof of Theorem 1.3 is established in Section 5.

### 3. GREEN'S REGULARITY LEMMA FOR $(b, \sigma)$ -SPARSE CAYLEY GRAPHS

In this section, unless otherwise specified, all Fourier transforms and convolutions are taken with respect to an underlying subspace  $H$ . For short, we let  $N = |V|$ .

**Theorem 3.1.** *Let  $\alpha, \varepsilon \in (0, 1)$ . There is a constant  $\sigma = \sigma(\varepsilon, \alpha)$  such that the following holds. Let  $R$  be a  $\sigma$ -regular set of  $V$  and let  $A$  be a subset of  $R$  of cardinality  $\alpha|R|$ . Then there is a subspace  $H \leq V$  of index at most  $W(4(\varepsilon\alpha)^{-2})$  which is  $\varepsilon$ -regular for  $A$ .*

We pause to discuss the strength of Theorem 3.1. First, since  $R$  is  $\sigma$ -regular, the Cayley graph  $G_A$  generated by  $A$  is  $(2/\alpha, \sigma)$ -sparse. Indeed, for any  $X, Y \in V$  such that  $|X| \geq \sigma|V|$  and  $|Y| \geq |V|$  we have  $e_{G_A}(X, Y) \leq e_{G_R}(X, Y)$ . On the other hand, since  $R$  is  $\sigma$ -regular, we have

$$\begin{aligned} e_{G_R}(X, Y) &= (1 + o_\sigma(1))|R||X||Y|/N \\ &\leq 2|R||X||Y|/N \\ &\leq (2/\alpha)|A||X||Y|/N \\ &= (2/\alpha)|X||Y|d(G_A). \end{aligned}$$

Hence,  $d_{G_A}(X, Y) \leq (2/\alpha)d(G_A)$ .

Now, since  $G_A$  is  $(2/\alpha, \sigma)$ -sparse, Theorem 1.2 is applicable to  $G_A$ . The advantage of Theorem 3.1 is, besides implying Theorem 1.2, it also provides a well-structured partition for the vertex set of  $G_A$  as follows.

Let  $V = \cup_{i=1}^K H_i$  be the partition of  $V$  into affine translates of  $H$ . Let  $v_1, \dots, v_K$  be representatives of the coset subgroups  $V/H$ . Then by definition, all but at most  $\varepsilon K$  vectors  $v_1, \dots, v_K$  are  $\varepsilon$ -regular vectors with respect to  $H$ .

Next assume that  $H_i = v_i + H$  and  $H_j = v_j + H$  are two affine translates of  $H$  such that  $v_j - v_i$  is an  $\varepsilon$ -regular vector. We will show that the subgraph  $G_A(H_i, H_j)$  is relatively  $\varepsilon^{1/3}$ -regular.

It is clear that  $e_{G_A}(H_i, H_j) = |H||A_H^{v_j - v_i}|$ ; thus

$$d_{G_A}(H_i, H_j) = |A_H^{v_j - v_i}|/|H|.$$

Let  $X \subset H_i$  and  $Y \subset H_j$  be any two subsets of  $H_i$  and  $H_j$  respectively, which satisfy  $|X|, |Y| \geq \varepsilon^{1/3}|H|$ . We shall estimate the number of edges generated by  $X$  and  $Y$ . We have

$$\begin{aligned}
e_{G_A}(X, Y) &= \sum_{x \in H_i, y \in H_j} A(y - x)X(x)Y(y) \\
&= \sum_{x', y' \in H} A_H^{v_j - v_i}(y' - x')X(x' + v_i)Y(y' + v_j) \\
&= \sum_{x', y' \in H} A_H^{v_j - v_i}(y' - x')(X - v_i)(x')(Y - v_j)(y')
\end{aligned}$$

Now we apply the Fourier inversion formula to the last sum,

$$\begin{aligned}
e_{G_A}(X, Y) &= |H|^2 \sum_{\xi \in H} \widehat{A_H^{v_j - v_i}}(\xi) \widehat{(X - v_i)}(-\xi) \widehat{(Y - v_j)}(\xi) \\
&= |A_H^{v_j - v_i}| |X| |Y| / H + \sum_{\xi \in H \setminus \{0\}} \widehat{A_H^{v_j - v_i}}(\xi) \widehat{(X - v_i)}(-\xi) \widehat{(Y - v_j)}(\xi).
\end{aligned}$$

Since  $v_j - v_i$  is an  $\varepsilon$ -regular vector with respect to  $H$ , we infer that

$$\left| e_{G_A}(X, Y) - |A_H^{v_j - v_i}| |X| |Y| / H \right| \leq (\varepsilon |A^{v_j - v_i}| / N) \sum_{\xi} \left| \widehat{(X - v_i)}(-\xi) \widehat{(Y - v_j)}(\xi) \right|.$$

By Parseval's identity and by the Cauchy-Schwarz inequality we thus have

$$\begin{aligned}
\left| e_{G_A}(X, Y) - |A_H^{v_j - v_i}| |X| |Y| / H \right| &\leq |H| (\varepsilon |A^{v_j - v_i}| / N) (|X| |Y|)^{1/2} \\
&\leq \varepsilon |A^{v_j - v_i}| |H|^2 / N.
\end{aligned}$$

It follows that

$$\begin{aligned}
|d_{G_A}(X, Y) - d_{G_A}(H_i, H_j)| &\leq \varepsilon |A^{v_j - v_i}| |H|^2 / (|X| |Y| N) \\
&\leq \varepsilon^{1/3} |A^{v_j - v_i}| / N \\
&= \varepsilon^{1/3} d(G_A).
\end{aligned}$$

Hence  $G_A(H_i, H_j)$  is indeed relatively  $\varepsilon^{1/3}$ -regular. One observes that  $v_j - v_i$  is an  $\varepsilon$ -regular vector for all but at most  $\varepsilon K^2$  pairs  $(i, j)$ . Hence the partition  $V = \cup_{i=1}^K H_i$  satisfies the conclusion of Theorem 1.2.

Another crucial observation, which will be used later on in applications, is that the definition of  $\varepsilon$ -regular vector works for any type of (linear) Cayley graph. For instance assume that  $(v_1 + v_2)/2$  is an  $\varepsilon$ -regular vector with respect to  $H$  and

define a bipartite Cayley graph  $G'_A$  on  $(H - v_1, H - v_2)$  by connecting  $(h_1 - v_1)$  with  $(h_2 - v_2)$  if  $((h_1 - v_1) + (h_2 - v_2))/2 = (h_1 + h_2) - (v_1 + v_2)/2 \in A$ . Then this graph  $G'_A$  is also  $\varepsilon^{1/3}$ -regular. To justify this fact, the reader just needs to follow the same lines of verification used for  $G_A$  above.

Now we start to prove Theorem 3.1.

**Proof .** Define  $d(A, H)$  by

$$d(A, H) := \frac{1}{N} \sum_{v \in V} \left( \frac{|A_H^v|}{|H|} \right)^2 / \left( \frac{|A|}{N} \right)^2.$$

Observe that  $d(A, H)$  is the mean of the squares of the normalized densities of the  $G_A(H_i, H_j)$ 's. We show that this quantity is always bounded.

**Claim 3.2.** *We have  $d(A, H) \leq 4/\alpha^2$  for any  $|H| \geq \sigma N$ .*

**Proof** (of Claim 3.2). Since  $H \geq \sigma N$ , by the  $\sigma$ -regularity of  $R$ , for any  $v$  we have,

$$|H||R_H^v| = e_{G_R}(H, H - v) = (1 + o_\sigma(1))|H||H||R|/N.$$

Hence  $|A_H^v|/|H| \leq |R_H^v|/|H| \leq 2|R|/N \leq (2/\alpha)|A|/N$ . As a result,

$$d(A, H) \leq \frac{1}{N} \sum_{v \in V} (2/\alpha)^2 \leq 4/\alpha^2.$$

■

As in the proof of Szemerédi's regularity lemma, when a partition with too many irregular pairs comes into play, then we pass to a finer partition, and by so the mean square of the densities will increase. What we are going to do is similar, the only difference is we restrict ourselves to a special family of partitions.

**Lemma 3.3.** *Let  $\varepsilon \in (0, 1)$  and suppose that  $H$  is a subspace of  $V$ , which is not  $\varepsilon$ -regular for  $A$ . Then there is a subspace  $H' \leq H$  such that  $|V/H'| \leq (2p)^{|G/H|}$  and  $d(A, H') \geq d(A, H) + \varepsilon^3$ .*

**Proof** (of Lemma 3.3). Since  $H$  is not  $\varepsilon$ -regular for  $A$ , there are  $\varepsilon N$  vectors  $v$  such that  $\sup_{\xi \notin H^\perp} |\widehat{A_H^v}(\xi)| \geq \varepsilon|A||H|/N$ . In other words, there exists a positive integer  $m$  satisfying  $\varepsilon N/|H| \leq m \leq N/|H|$  together with  $m$  coset representatives  $v_1, \dots, v_m \in V/H$  and vectors  $\xi_1, \dots, \xi_m \in H$ , where  $N/|H| \geq m \geq \varepsilon N/|H|$ , such that

$$|\widehat{A_H^{v_i}}(\xi_i)| \geq \varepsilon|A|/N.$$

Now let  $H' \subset H$  be the annihilator of all  $\xi_i$ 's. It is clear that

$$|H'| \geq |H|/p^m \geq |H|/p^{|V/H|}$$

Hence,

$$|V/H'| \leq |V/H|p^{|V/H|} < (2p)^{|V/H|}.$$

Set  $S := N|H'|^2(|A|/N)^2|H|d(A, H')$ . It is obvious that

$$S = |H| \sum_{v \in V} |A_{H'}^v|^2 = \sum_{v \in V, h \in H} |A_{H'}^{v+h}|^2.$$

Notice that  $|A_{H'}^{v+h}| = \sum_{x \in H} (A+v)(x-h)H'(x) = \sum_{x \in H} (A+v)(x)H'(x+h) = |H|(A_H^v * H')(-h)$ . We rewrite  $S$  and then use Plancherel's formula,

$$\begin{aligned} S &= |H|^2 \sum_{v \in V, h \in H} |A_H^v * H'(h)|^2 \\ &= |H|^3 \sum_{v \in V, \xi \in H} \left| \widehat{A_H^v * H'}(\xi) \right|^2 \\ &= |H|^3 \sum_{v \in V, \xi \in H} |\widehat{A_H^v}(\xi)|^2 |\widehat{H'}(\xi)|^2. \end{aligned}$$

In the last sum, the contribution of the  $\xi = 0$  term gives

$$\begin{aligned} S_0 &= |H|^3 \sum_{v \in V} (|A_H^v|/|H|)^2 (|H'|/|H|)^2 \\ &= |H||H'|^2 \sum_{v \in V} (|A_H^v|/|H|)^2 \\ &= N|H||H'|^2(|A|/N)^2 d(A, H); \end{aligned}$$

while the sums contributed from  $\xi \in H \setminus \{0\}$  is bounded from below by

$$S_{\neq 0} \geq |H|^3 \sum_{i=1}^m \sum_{v \in H+v_i} |\widehat{A_H^{v_i}}(\xi_i)|^2 |\widehat{H'}(\xi_i)|^2.$$

But since  $\xi_i \in H'^\perp$ , we have  $\widehat{H'}(\xi_i) = |H'|/|H|$ . Use the bound  $|\widehat{A_H^{v_i}}(\xi_i)| \geq \varepsilon|A|/N$  for all  $1 \leq i \leq m$ , we obtain



$$\begin{aligned}
S_{\neq 0} &\geq |H|^3 m |H| (\varepsilon |A|/N)^2 (|H'|/|H|)^2 \\
&\geq |H|^3 (\varepsilon |N|/|H|) |H| \varepsilon^2 (|A|/N)^2 (|H'|/|H|)^2 \\
&= \varepsilon^3 |H| |H'|^2 N (|A|/N)^2.
\end{aligned}$$

From the estimate for  $S_0$  and  $S_{\neq 0}$  we conclude that  $d(A, H') \geq d(A, H) + \varepsilon^3$ . ■

To complete the proof of Theorem 3.1 we keep applying Lemma 3.3. Since  $d(A, H) \leq 4/\alpha^2$ , the iteration stops after at most  $4\varepsilon^{-3}\alpha^{-2}$  steps. During the iteration,  $|H'|$  is always bounded below by  $N/W(4\varepsilon^{-3}\alpha^{-2})$ , thus we may choose  $\sigma = (2W(4\varepsilon^{-3}\alpha^{-2}))^{-1}$ . ■

Let us conclude this section by mentioning an important corollary of Theorem 3.1.

**Theorem 3.4.** *Let  $\alpha, \varepsilon \in (0, 1)$  and let  $m$  be a positive integer. There is a constant  $\sigma = \sigma(\varepsilon, \alpha, m)$  such that if  $R$  is a  $\sigma$ -regular set of  $V$  and  $A$  is a subset of  $R$  of cardinality  $\alpha|R|$ , then the following holds. Assume that  $A = \cup_{i=1}^m A_i$  is a partition of  $A$  into  $m$  distinct sets of size  $|A|/m$ . Then there is a subspace  $H \leq V$  of index bounded by  $W(4m^2\varepsilon^{-3}\alpha^{-2})$  which is  $\varepsilon$ -regular for all  $A_i$ 's.*

To prove Theorem 3.4 first we let  $d(A_1, \dots, A_m, H) := \sum_{i=1}^m d(A_i, H)$ . Next keep iterating Lemma 3.3 if  $H$  is not  $\varepsilon$ -regular for some  $A_i$ . Since  $d(A_1, \dots, A_i, H) \leq 4m^2/\alpha^2$ , the iteration will stop after at most  $4m^2\varepsilon^{-3}\alpha^{-2}$  steps.

## 4. MAIN LEMMAS FOR APPLICATIONS

### 4.1. Regularity of a random set.

**Lemma 4.2.** *For  $\sigma > 0$  there is a constant  $C(\sigma)$  such that if  $r \geq C(\sigma)N^{1/2}$  and  $R$  is a random subset of size  $r$  of  $V$ , then  $R$  is a  $\sigma$ -regular set almost surely.*

To start with, we consider a slightly different model as follows.

**Lemma 4.3.** *For  $\sigma > 0$  there is a constant  $C(\sigma)$  such that if  $r \geq C(\sigma)N^{1/2}$  and  $q = r/N$ , and  $R$  is a subset of  $V$  whose elements are equally selected with probability  $q$ , then  $R$  is a  $\sigma$ -regular set almost surely.*

**Proof** (of Lemma 4.3) Let  $X, Y \subset V$ , of cardinality at least  $\sigma N$ . The number of edges of  $G_R$  generated by  $X$  and  $Y$  is

$$e_R(X, Y) = \sum_{x, y \in V} 1_R(y - x) 1_X(x) 1_Y(y) = N^2 \sum_{\xi \in V} \widehat{1_R}(\xi) \widehat{1_X}(\xi) \widehat{1_Y}(-\xi)$$

where the Fourier transform is defined with respect to  $V$ , and the latter identity comes from Fourier inversion formula. Thus we have

$$e_R(X, Y) = |R||X||Y|/N + N^2 \sum_{\xi \in V, \xi \neq 0} \widehat{1}_R(\xi) \widehat{1}_X(\xi) \widehat{1}_Y(-\xi).$$

Let us pause to estimate  $\widehat{1}_R(\xi)$ .

**Lemma 4.4.**  $\sup_{\xi \neq 0} |\widehat{1}_R(\xi)| < |R|/(N \log N)$  *almost surely for  $R$ .*

The proof of this lemma is routine by applying the exponential moment method. For the sake of completeness, we prove it in Appendix A.

Assuming Lemma 4.4, then by the Cauchy-Schwarz inequality and Parseval's identity we have

$$\begin{aligned} |e_R(X, Y) - |R||X||Y|/N| &\leq N^2 \sup_{\xi \neq 0} |\widehat{1}_R(\xi)| \left( \sum_{\xi \in V} |\widehat{1}_X(\xi)|^2 \sum_{\xi \in V} |\widehat{1}_Y(\xi)|^2 \right)^{1/2} \\ &\leq N^2 \sup_{\xi \neq 0} |\widehat{1}_R(\xi)| (|X||Y|/N^2)^{1/2} \\ &= \sup_{\xi \neq 0} |\widehat{1}_R(\xi)| (|X||Y|)^{1/2} N. \end{aligned}$$

On the other hand, as  $|X|, |Y| \geq \sigma N/4$  and  $\sup_{\xi \neq 0} |\widehat{1}_R(\xi)| \leq |R|/(N \log N)$ , we have

$$\sup_{\xi \neq 0} |\widehat{1}_R(\xi)| (|X||Y|)^{1/2} = o(|R||X||Y|/N),$$

completing the proof of Lemma 4.3. ■

Next we show that the two models, of Lemma 4.2 and of Lemma 4.3, are similar.

**Proof** (of Lemma 4.2). Let  $q = (1 - \sigma^4)|R|/N$ . We first consider a random set  $R_1$  by selecting each element of  $V$  with probability  $q$ . It is obvious that the size of this random set belongs to  $[(1 - 2\sigma^4)|R|, |R|]$  almost surely. We restrict ourself to this event by renormalizing the probability space. Hence the random set  $R_1$  is chosen uniformly from the collection of subsets of size  $[(1 - 2\sigma^4)|R|, |R|]$ . Next we pick uniformly a set  $R_2$  of size  $|R| - |R_1|$  from  $V \setminus R_1$  and set  $R = R_1 \cup R_2$ .

Suppose that  $X, Y \subset V$  and  $|X|, |Y| \geq \sigma N$ . Since  $R_1$  is  $\sigma$ -regular almost surely by Lemma 4.3, we have  $(1 - o_\sigma(1))|R_1||X||Y|/N \leq e_{R_1}(X, Y) \leq (1 + o_\sigma(1))|R_1||X||Y|/N$ . On the other hand, it is obvious that

$$\begin{aligned} e_{R_1}(X, Y) &\leq e_R(X, Y) \\ &\leq e_{R_1}(X, Y) + |R_2|N \\ &\leq (1 + o_\sigma(1))|R||X||Y|/N + 2\sigma^4|R|N \\ &= (1 + o_\sigma(1))|R||X||Y|/N. \end{aligned}$$

Hence  $e_R(X, Y) = (1 + o_\sigma(1))|X||Y|/N$  almost surely, completing the proof of Lemma 4.2.  $\blacksquare$

**4.5. Edge distribution of quasi-random graphs.** Roughly speaking, if we choose randomly a large number of vertices of a dense quasi-random graph, then the chance of obtaining an edge is very high. This simple observation, as a strong tool to exploit structure for counting, was used in [2], and will play a key role in our proof of Theorem 1.3.

Let  $G = G(u, \rho, \varepsilon)$  be an  $\varepsilon$ -regular bipartite graph,  $V(G) = U_1 \cup U_2$ , where  $|U_1| = |U_2| = u$  and  $d(G) = e(G)/u^2 = \rho$ . Let  $t_1, t_2 < u/2$  be two given positive integers. We select a random subgraph of  $G$  as follows. First, an adversary chooses a set  $S_1 \subset U_1$  with  $|S_1| \leq u/2$ . Then we pick a set  $T_1 \subset U_1 \setminus S_1$  with  $|T_1| = t_1$  from the collections of all  $d_1$ -subsets of  $U_1 \setminus S_1$  with equal probability. Next, our adversary picks a set  $S_2 \subset U_2$  with  $|S_2| \leq u/2$ , and we pick a set  $T_2 \subset U_2 \setminus S_2$  with  $|T_2| = t_2$  from the collections of all  $d_2$ -subsets of  $U_2 \setminus S_2$  with equal probability. Let us call the outcome of the above procedure a random  $(t_1, t_2)$ -subgraph of  $H$ .

**Lemma 4.6.** [2, Lemma 11] *For every constant  $0 < \eta < 1$ , there exist a constant  $0 < \varepsilon < 1$  and a natural number  $u_0$  such that, for any real  $t \geq 2(u/\varepsilon)^{1/2}$  and any given graph  $G = G(u, \rho, \varepsilon)$  as above with  $u \geq u_0$  and  $\rho \geq t/u$ , the following holds. If  $t_1, t_2 \geq t$ , regardless of the choices for  $S_1$  and  $S_2$  of our adversary, the probability that a random  $(t_1, t_2)$ -subgraph of  $G$  fails to contain an edge is at most  $\eta^t$ .*

The proof of Lemma 4.6 is simple, the interested reader may read [2].

**4.7. Roth's theorem for  $\mathbb{F}_p^n$ .** Another important ingredient which will serve as a starting point for our argument is Roth's theorem.

**Theorem 4.8.** *For any  $\delta > 0$  there is a number  $c(\delta) > 0$  such that if  $B$  is a subset of  $V$  of size  $\delta|V|$ , then  $B$  contains at least  $c(\delta)|V|^2$  three-term arithmetic progressions.*

In the next section, we shall put every thing together to establish Theorem 1.3.

## 5. PROOF OF THEOREM 1.3

First, by Theorem 4.8, it is enough to work with the case

$$r = o_\alpha(N).$$

We say that a set  $A$  which belongs to some  $\sigma$ -regular set  $R$  is  $(\alpha, \sigma)$ -bad if  $|A| = \alpha|R| = \alpha r$  and it contains no nontrivial three-term arithmetic progression. Our main goal is to give an upper bound for the number of bad sets of a given size.

**Theorem 5.1.** *Let  $\alpha$  and  $\eta$  be given positive numbers. Then there exist constants  $c = c(\alpha) > 0$ ,  $C = C(\eta, \alpha) > 0$  and  $\sigma = \sigma(\alpha, \eta) > 0$  such that for all  $s \geq C(\alpha, \eta)N^{1/2}$ , the number of  $(\alpha, \sigma)$ -bad sets of size  $s$  is at most  $\eta^{c(\alpha)s} \binom{n}{s}$ .*

**Proof** (of Theorem 1.3 assuming Theorem 5.1). We choose  $\eta = \eta(\alpha)$  to be small enough. Let  $s \geq C(\alpha, \eta)N^{1/2}$  and put  $r = s/\alpha$ . Pick a random set  $R$  among all  $r$ -subsets of  $[n]$ . Then by Theorem 4.2,  $R$  is  $\sigma$ -regular almost surely. Among these  $\sigma$ -regular  $r$ -sets, by Theorem 5.1, the number of sets that contain at least an  $(\alpha, \sigma)$ -bad subset is at most

$$\eta^{c(\alpha)s} \binom{n}{s} \binom{n-s}{r-s}.$$

Observe that, as  $\eta$  is small enough, this amount is  $o\left(\binom{n}{r}\right)$ . Hence almost all  $r$ -sets of  $[n]$  contain no bad subsets at all. To finish the proof, we note that if  $R$  contains no  $(\alpha, \sigma)$ -bad subset, then it is  $(\alpha, 3AP)$ -dense. ■

We shall concentrate on proving Theorem 5.1 by localizing some properties of  $A$ . Our approach follows that of [2] closely, but the key difference here is that we shall exploit rich structure obtained from Theorem 3.1 and Theorem 3.4.

Let  $R$  be a  $\sigma$ -regular of fixed size  $C(\sigma)N^{1/2} \leq r = o(N)$  such that  $A \subset R$ . Let  $m = m(\alpha)$  be a large number to be defined later.

From now on we shall view  $A$  as an ordered  $m$ -set-tuple,  $A = (A_1, \dots, A_m)$  where  $|A_i| = |A|/m$  for all  $i$  and  $A = \cup A_i$ . We shall choose  $\varepsilon = \varepsilon(\alpha)$  to be small enough. By Theorem 3.4, there exists a subspace  $H$  of  $V$  which has index bounded by  $W(4m^2\alpha^{-2}\varepsilon^{-3})$  and which is  $\varepsilon$ -regular for all  $A_i$ 's.

Let  $v_1, \dots, v_K$  be representatives of the coset subgroup  $V' := V/H$ . For each  $A_i$ , let us consider a set  $B_i$  of vectors  $v$  that satisfy the following conditions:

- $v$  is  $\varepsilon$ -regular with respect to  $A_i$  and  $H$ .

- $|(A_i)_H^v| \geq (1/4)|A_i||H|/N$ .

It is clear that  $|(A_i)_H^v| \leq A_H^v \leq |R_H^v|$ . But by definition of  $R$ ,  $|R_H^v| \leq 2|R||H|/N$ ; thus we have

$$\begin{aligned} \sum_{v \in B_i} |(A_i)_H^v| &\geq |A_i| - (\varepsilon K)(2|R||H|/N) - K((1/4)|A_i||H|/N) \\ &\geq (1 - (\varepsilon m)/\alpha - 1/4)|A_i| \geq |A_i|/2, \end{aligned}$$

provided that  $\varepsilon \leq \alpha/2m$ . We infer that the size of  $B_i$  is large,

$$|B_i| \geq (|A_i|/2)/(2|R||H|/N) \geq \frac{\alpha}{4m}K.$$

By a truncation if needed, we assume that  $B_i$  has cardinality  $(\alpha/4m)K$  for all  $i$ . Notice that these sets are not necessarily disjoint. We shall show that there are many three-term arithmetic progressions (in  $V'$ ) with the property that all 3 terms belong to different  $B_i$ 's.

Now we set  $B := \{v \in V' : v \in B_i \cap B_j \cap B_k \text{ for some } i < j < k\}$  and consider the following two cases.

**Case 1.**  $|B| \geq (\alpha/8m)K = (\alpha/8m)|V'|$ . Applying Theorem 4.8 we obtain  $c(\alpha/8m)K^2$  three-term arithmetic progressions in  $B$ . By the definition of  $B$ , it follows that there are  $c(\alpha/8m)K^2$  three-term arithmetic progressions with the property that all three terms belong to three different sets  $B_i$ .

**Case 2.**  $|B| \leq (\alpha/8m)K = |B_i|/2$ . We let  $B' = \cup_{i=1}^m B_i \setminus B$ . By an elementary counting argument, it follows that  $|B'| \geq m|B_i|/4 = (\alpha/16)K$ . Let us write  $B' = \cup_{i=1}^m B'_i$ , where  $B'_i \subset B_i$  and all  $B'_i$  are disjoint.

By Theorem 4.8, the set  $B'$  contains  $c(\alpha/16)K^2$  three-term arithmetic progressions. Among them, since each three-term arithmetic progression is defined by two parameters, the number of three-term arithmetic progressions that consist of at least two terms from the same  $B'_i$  is bounded by  $3 \sum_{i=1}^m |B'_i|^2$ . The latter quantity is bounded by  $3|B_i|(\sum_{i=1}^m |B_i|) \leq 3(\alpha/4m)(\alpha/4)K^2$ ; which is negligible compared to  $c(\alpha/16)K^2$  by letting  $m = m(\alpha)$  large.

In both cases, the number of three-term arithmetic progressions with the property that all three terms belong to three different sets  $B_i$  is at least  $c'(\alpha)K^2$ , where  $c'(\alpha) = \min(c(\alpha/8m), c(\alpha/16)/2)$ . By an averaging argument, there exist three indices  $i_0 < j_0 < k_0$  such that the number of three-term arithmetic progressions in  $B_{i_0} \times B_{j_0} \times B_{k_0}$  is at least  $c'(\alpha)K^2/m^2 = c''(\alpha)K^2$ . In particular, there exist a vector  $v_{i_0} \in B_{i_0}$  and  $c''(\alpha)K$  pairs  $(v_{j_0}^l, v_{k_0}^l) \in B_{j_0} \times B_{k_0}$  such that each triple  $(v_{i_0}, v_{j_0}^l, v_{k_0}^l)$  is a three-term arithmetic progression.

Let us summarize what have been achieved.

- (1) There exists a subspace  $H$  of index bounded by a function of  $\alpha$  and  $\varepsilon$ , and there exist  $A_{i_0}, A_{j_0}, A_{k_0}$  and triples  $(v_{i_0}, v_{j_0}^l, v_{k_0}^l)$ , where  $1 \leq l \leq c'''(\alpha)K$ , such that the following holds:
- (2)  $v_{i_0}$  is an  $\varepsilon$ -regular vector for  $A_{i_0}$ , and  $(A_{i_0})_H^{v_{i_0}} \geq (1/4)|A_{i_0}||H|/N$ ;
- (3)  $(A_{j_0})_H^{v_{j_0}^l} \geq (1/4)|A_{j_0}||H|/N = (1/4m)s|H|/N$ ;
- (4)  $(A_{k_0})_H^{v_{k_0}^l} \geq (1/4)|A_{k_0}||H|/|V| = (1/4m)s|H|/N$ ;
- (5)  $(v_{i_0}, v_{j_0}^l, v_{k_0}^l)$  is a three-term arithmetic progression in  $V/H$ .

One also observes that  $v_{j_0}^l, v_{k_0}^l$  are  $\varepsilon$ -regular vectors with respect to  $A_{j_0}$  and  $A_{k_0}$ ; but we do not need this fact. Since this configuration arises from [2], let us call it an  $(\alpha, \varepsilon, H, i_0, j_0, k_0, v_{i_0}, v_{j_0}^l, v_{k_0}^l)$ -flower. Roughly speaking, the reader may visualize a flower with a center  $A_{i_0} + v_{i_0} \cap H$ , where  $A_{i_0} + v_{i_0} \cap H$  sits nicely in  $H$ , and with  $c'''(\alpha)K$  petals  $(A_{j_0} + v_{j_0}^l \cap H, A_{k_0} + v_{k_0}^l \cap H)$ .

We denote by  $\mathcal{S}$  the collections of all ordered  $m$ -set-tuples  $A = (A_1, \dots, A_m)$  of size  $|A| = s$ .

**Proposition 5.2.** *Let  $\alpha, \eta$  be given. Then there exist constants  $c = c(\alpha) > 0, C = C(\alpha, \eta)$  and  $\varepsilon = \varepsilon(\alpha, \eta) > 0$  such that the number of ordered  $m$ -set-tuples  $A = (A_1, \dots, A_m)$  of size  $s$ , where  $s \geq C(\alpha, \eta)N^{1/2}$ , that contain a flower but not any non-trivial three-term arithmetic progression is at most  $\eta^{c(\alpha)s}|\mathcal{S}|$ .*

It is clear that Proposition 5.2 implies Theorem 5.1. Hence we just need to prove Proposition 5.2.

First of all we shall estimate the probability that a set  $A$  that contains a given  $(\alpha, \varepsilon, i_0, j_0, k_0, v_{i_0}, v_{j_0}^l, v_{k_0}^l)$ -flower but contains no non-trivial three-term arithmetic progressions. On this probability space we also fix  $A_{i_0}$  and fix the size of  $A_{j_0}^l \cap H$  and  $A_{k_0}^l \cap H$  for all  $l$ . Hence,  $v_{i_0}$  is an  $\varepsilon$ -regular vector with respect to a fixed sets  $A_{i_0}$  and  $H$ ; and the sets  $A_{j_0}$  and  $A_{k_0}$  vary in such a way that  $v_{j_0}^l, v_{k_0}^l$  satisfy (3) and (4) respectively (in other words,  $A_{j_0}, A_{k_0}$  intersects  $H - v_{j_0}^l, H - v_{k_0}^l$  in sets of given size).

Without loss of generality, we assume that  $2v_{i_0} = v_{j_0}^l + v_{k_0}^l$  for all  $l$ . We define a Cayley graph between  $H - v_{j_0}^l$  and  $H - v_{k_0}^l$  by connecting  $v_1 \in H - v_{j_0}^l$  to  $v_2 \in H - v_{k_0}^l$  if  $(v_1 + v_2)/2 \in A_{i_0}$ . Since  $v_{i_0}$  is  $\varepsilon$ -regular with respect to  $A_{i_0}$ , by the observation made before proving Theorem 3.1, this graph is also  $\varepsilon$ -regular.

By choosing  $\varepsilon = \varepsilon(\alpha, \eta) = \varepsilon(\alpha)$  to be small enough, and recalling that  $N^{1/2} \ll s < r = o(N)$ , we may check that for each bipartite graph  $(H - v_{j_0}^l, H - v_{k_0}^l)$ , the

assumptions of Lemma 4.6 are satisfied with  $S_1 = \bigcup_{1 \leq m < j_0} (A_m \cap (H - v_{j_0}^l))$ ,  $T_1 = A_{j_0}^l \cap (H - v_{j_0}^l)$ ,  $S_2 = \bigcup_{1 \leq m < k_0} (A_m \cap (H - v_{k_0}^l))$ , and  $T_2 = A_{k_0} \cap (H - v_{k_0}^l)$ . It follows that the probability each petal fails to contain a three-term arithmetic progression is less than  $\eta^{(1/4)s/(mK)}$ . Hence the probability that  $A$  contains no non-trivial three-term arithmetic progressions is less than  $\eta^{(1/4)c''(\alpha)s/m} = \eta^{c'''(\alpha)s}$ .

Now we bound the number of flowers: the number of choices for  $H$  is bounded by  $N^{W(4m^2\alpha^{-1/2}\varepsilon^{-3})}$ , the number of choices for  $(i_0, j_0, k_0, v_{i_0}^l, v_{j_0}^l, v_{k_0}^l)$  is bounded by  $K^{4+2c''(\alpha)K}$  (which is independent of  $N$ ). Hence there are at most  $N^{C(\alpha)}$  flowers.

Putting everything together, we infer that the number of  $A$  that contains some flower but not any non-trivial three-term arithmetic progression is at most

$$N^{C(\alpha)} \eta^{c'''(\alpha)s} |\mathcal{S}| \leq \eta^{c''''(\alpha)s} |\mathcal{S}|,$$

completing the proof.

#### APPENDIX A. PROOF OF LEMMA 4.4

Without loss of generality, we just work with the real part of  $\widehat{1}_R$ . We shall prove  $\mathbf{P}_R(\sup_{\xi \neq 0} |\Re \widehat{1}_R(\xi)| \geq \lambda/N) = o(1)$  for some appropriately chosen  $\lambda$ . Since the treatment for other cases is similar, we just show that  $\mathbf{P}(\Re \widehat{1}_R(\xi) \geq \lambda/N)$  is very small for each fixed  $\xi \neq 0$ . For convenience, put

$$X = N \Re \widehat{1}_R(\xi) = \sum_{v \in V} 1_R(v) \Re e(-\langle v, \xi \rangle) := \sum_{v \in V} X_v.$$

One observes that  $X$  is a sum of  $N$  linearly independent real variables  $X_v$ 's. Choosing  $t$  to be a positive number smaller than 1, we have

$$\begin{aligned} \mathbf{P}_R(X \geq \lambda) &= \mathbf{P}_R(\exp(tX) \geq \exp(t\lambda)) \\ &\leq \mathbf{E}(\exp(tX)) / \exp(t\lambda) \\ &= \prod \mathbf{E}(\exp(tX_v)) / \exp(t\lambda) \\ &= \exp(t\mathbf{E}X) \prod \mathbf{E}(\exp(tX_v - t\mathbf{E}(X_v))) / \exp(t\lambda) \\ &= \exp(t\mathbf{E}X) \prod \mathbf{E}(\exp(tY_v)) / \exp(t\lambda), \end{aligned}$$

where  $Y_v := X_v - \mathbf{E}(X_v) = (1_R(v) - q) \Re e(-\langle v, \xi \rangle)$ .

Notice that  $|Y_v| \leq 1$  and  $0 < t \leq 1$ . We thus have  $\exp(tY_v) \leq 1 + tY_v + t^2Y_v^2$ . Hence

$$\mathbf{E}(\exp(tY_v)) \leq 1 + \mathbf{E}(t^2 Y_v^2) \leq \exp(\mathbf{E}(t^2 Y_v^2)).$$

Also, because  $\mathbf{E}X = q\Re \sum_{v \in V} e(-\langle v, \xi \rangle) = 0$ , it follows that

$$\mathbf{P}(X \geq \lambda) \leq \prod \mathbf{E}(\exp(tY_v))/\exp(t\lambda) \leq \exp(t^2 \sum_{v \in V} \mathbf{E}(Y_v^2))/\exp(t\lambda).$$

On the other hand, it is clear from the definition of  $Y_v$  that  $\sum_{v \in V} \mathbf{E}(Y_v^2) \leq qN$ . Thus

$$\mathbf{P}(X > \lambda) \leq \exp(t^2 qN - t\lambda).$$

By choosing  $\lambda = |R|/\log N$  and  $t = \lambda/(2qN) = 1/(2 \log N)$  (thus  $t < 1$ ), we deduce that

$$P(X \geq |R|/\log N) \leq \exp(-|R|/(2 \log^2 N)).$$

Hence

$$\mathbf{P}(\sup_{\xi \neq 0} \Re \widehat{1}_S(\xi) > |R|/(N \log N)) \leq N \exp(-|R|/(2 \log^2 N)) = o(1).$$

(Note that the choice for  $\lambda$  above is not optimal, but it is enough for our goal.)

**Acknowledgement.** The author would like to thank prof. Van H. Vu for discussions and encouragement. He is also grateful to Philip M. Wood for reading the note and giving many valuable suggestions.

#### REFERENCES

- [1] **B. Green**, *A Szemerédi-type regularity lemma in A belian groups, with applications*, GAFA 15 (2005), no. 2, 340-376.
- [2] **Y. Kohayakawa**, **T. Łuczak** and **V. Rödl**, *Arithmetic progressions of length three in subsets on a random sets*, Acta Arith. 75 (1996), 133-163.
- [3] **H. Nguyen**, *On two-point additive configurations in random sets*, Integers 9 (2009), 41-45.
- [4] **T. Tao** and **V. Vu**, *Additive Combinatorics*, Cambridge University Press, 2006.

DEPARTMENT OF MATHEMATICS, RUTGERS UNIVERSITY, PISCATAWAY, NJ 08854, USA